

Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) **EP 0 814 398 A1**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
**29.12.1997 Bulletin 1997/52**

(51) Int Cl.<sup>6</sup>: **G06F 1/00**

(21) Application number: **97303552.0**

(22) Date of filing: **23.05.1997**

(84) Designated Contracting States:  
**GB**

(72) Inventor: **Akatsu, Masaharu**  
**D202 Menlo Park, CA 94025 (US)**

(30) Priority: **24.05.1996 JP 129572/96**

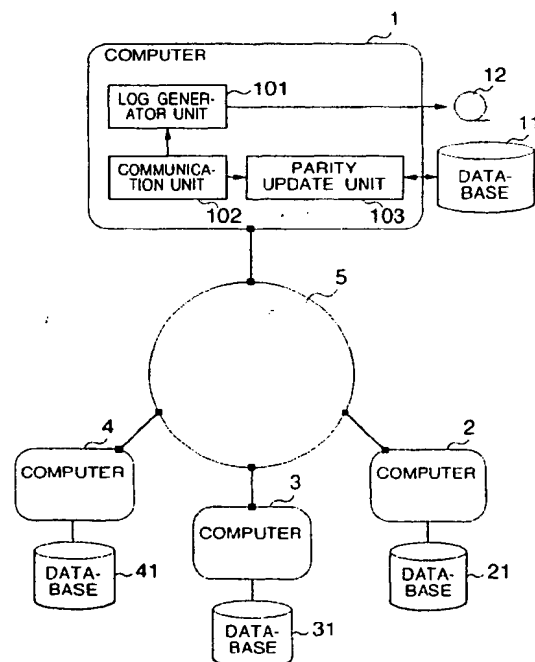
(74) Representative: **Hackney, Nigel John et al**  
**Mewburn Ellis,**  
**York House,**  
**23 Kingsway**  
**London WC2B 6HP (GB)**

(71) Applicant: **HITACHI, LTD.**  
**Chiyoda-ku, Tokyo 101 (JP)**

(54) **Method and system for detecting fraudulent data update**

(57) A method and system for detecting fraudulent data update of databases (21, 31, 41) connected to a plurality of distributed computers (2, 3, 4) in which a monitor computer (1) connected via a network (5) to the distributed computers detects fraudulent data update. The monitor computer (1) collects initial data of the databases (21, 31, 41) of the distributed computers via the network (5) to generate parities for data at same storage fields and store them in a monitor database (11). Each time the databases (21, 31, 41) connected to the distributed computers are updated, a parity update unit (103) of the monitor computer generates new parities from data before and after the update and old parities and replace the old parities stored in the monitor database (11) by the new parities. At an auditing time, the monitor computer (1) collects latest data of the distributed databases (21, 31, 41), and generates parities which are compared with the corresponding parities already stored in the monitor database (11) to detect an inconsistency of both the parities.

**FIG. 1**



**EP 0 814 398 A1**

## Description

The present invention relates to updating distributed databases, and more particularly to a method and system for detecting fraudulent data update. Description of the Related Art

Operating systems (OS) of computers are provided with a security function such as access control and file protection in order to prevent fraudulent update of files and databases. For example, as a user logs in a computer, the operating system checks the password and a database access privilege to thereby limit database update to specific persons. Since the operation system stores data update log, it is possible to check fraudulent data update by unprivileged persons. Computer security technology is described, for example, in "Information Security", Nikkei McGraw-Hill Corp., 1987, pp. 243 - 249.

Conventional security technology is effective for the prevention of fraudulent data update by unprivileged persons, however, it cannot deal with fraud by insiders having data update privilege. It is not easy to detect fraudulent data update by insiders of databases, particularly databases distributed at factories, sales offices, business offices and the like of companies and monetary institutes.

Preferably, it is therefore an object of the present invention to provide a method and system for detecting fraudulent or unauthorized data update by insiders of databases of a distributed computer system, capable of allowing third parties to check fraud.

In order to achieve the above object, according to one aspect of the present invention, there is provided a method of detecting fraudulent or unauthorized data update for a computer system including distributed databases and local computers at local sales offices and local business offices, a monitor computer, and a network interconnecting the local and monitor computers, comprising the steps, to be executed by the monitor computer, of: generating parity data of initial data collected from respective sites of the databases at each of same storage fields and storing the generated parity data; generating, each time data in each database is updated, new parity data from data before and after the update and old parity data to replace the old parity data by the new parity data; comparing parity data generated at an auditing time from latest data stored in the databases at each of the same storage fields, with the parity data already stored; and determining, if the comparison result indicates an inconsistency of both the parity data, that data in the databases was updated fraudulently.

If the data in the distributed databases is not coincident with the corresponding data already transmitted to the monitor computer, both the parity data are inconsistent so that fraudulent data update can be detected.

Other objects, features and advantages of the present invention will become apparent from reading of the following description of the preferred embodiments

taken in conjunction with the accompanying drawings.

Fig. 1 is a schematic diagram showing the structure of a computer system according to an embodiment of the invention.

Fig. 2 is a diagram showing the data format of a database of transaction data of the embodiment.

Fig. 3 is a diagram illustrating a process of generating a parity from data of one data item.

Fig. 4 is a schematic diagram showing the structure of a computer system according to another embodiment of the invention.

Fig. 5A is a flow chart illustrating the procedures to be executed by computers of the embodiment.

Figs. 5B to 5D are flow charts illustrating the process steps (programs) to be executed by a distributed system including branch, server, and center (main business office) computers.

Fig. 6 is a flow chart illustrating the procedures to be executed by computers of the embodiment at an auditing time.

Embodiments of the invention will be described in detail with reference to the accompanying drawings.

Fig. 1 is a schematic diagram showing the structure of a computer system which can be used for embodying a method of detecting fraudulent data manipulation. Reference numerals 2, 3 and 4 represent computers installed at local business offices (branches) for processing information such as totaling sales amounts and customer management at each branch. Reference numerals 21, 31 and 41 represent databases for storing sales data and transaction data, the databases 23, 31 and 41 being stored in storage units connected to the computers 2, 3 and 4. Reference numeral represents a computer installed at a main business office (center) and connected to the computers 2, 3 and 4 via a network such as LAN. The computer 1 stores data before and after update received from the computers 2, 3 and 4 in its update data log, and calculates a new parity of the stored data to replace the already stored parity by the new parity. Reference numeral 12 represents the update log, and reference numeral 11 represents a relational database for storing parity information. The database 11 and update data log 12 are stored in a storage unit connected to the computer 1. Reference numeral 101 represents a log generator unit for generating log information in accordance with data received from the computers 2, 3 and 4 and storing it in the update data log 12. Reference numeral 103 represents a parity update unit for generating a new parity in accordance with data before and after update received from the computers 2, 3 and 4 and a corresponding old parity and replacing the old parity stored in the data base 11 by the new parity. Generating a new parity may be achieved through an EX-OR calculation between the old parity and a difference between data before and after update. Reference numeral 102 represents a communication unit for controlling information transfer via the network 5 to and from the computers 2, 3 and 4. The log gener-

ator unit 101 and parity update unit 103 are realized by executing programs stored in the storage unit of the computer 1. The communication unit 102 is realized by communication hardware of the computer 1 and execution of programs stored in the storage unit of the computer 1.

Fig. 2 shows an example of the format of business transaction data stored in one of databases of each of the databases 21, 31 and 41. The transaction data is constituted of a plurality of rows or records. Each row is constituted of data items including a transaction date (month/day), a customer number, a good number, a transaction quantity, a transaction money amount, and the like. The number of digits representing data of each data item takes a predetermined number. Deletion and addition of a row or record of this relational database is one kind of record update operations. Namely, deleting a record is an operation of changing the contents of an original record to all binary 0's, whereas adding a record is an operation of adding a record having the contents of all 0's.

Fig. 3 is a diagram illustrating the generation of a parity from data of each data item. For example, the data of transaction date (month/day) at the first row of transaction data of one of the relational databases 21, 31 and 41, shown at the left side of Fig. 3, is represented by binary values shown at the right side thereof, and the even parity generated from these data is shown at the right bottom thereof. Parities for other data items such as customer number are generated in the same manner as above by using data of customer numbers stored in the same storage field of the database of transaction data. The number of digits of a parity for each data item is the same as the number of digits of data in each data item.

Records of the update data log 12 are constituted of a transmission branch name, a reception time, a database name, a row number, new data and old data. The database 11 stores parity information in correspondence with each database in each of the databases 21, 31 and 41. For example, in the case of a database of transaction data, parity information is provided in correspondence with each data item such as date, customer number, and so on, and the total number of records is equal to the maximum number of transaction data in the databases 21, 31 and 41.

Prior to starting the system operation, the parity update unit 103 of the computer 1 collects as initial data the data stored in each database of each of the databases 21, 31 and 41 of the computers 2, 3 and 4, generates a parity for each data item of each record of each database, and stores it in the database 11. If there is a branch having no record in its database, no parity is generated, and parities for data items are generated only for those branches having records in their databases.

After the system starts, the computers 2, 3 and 4 at branches transmit new and old data to the computer 1 each time a new transaction occurs and record data is

updated. For addition/deletion of a record, an addition/deletion record with an addition/deletion discriminator is transmitted. The communication unit 102 of the computer 1 receives the transmitted data via the network 5, and the log generator unit 101 generates log data in accordance with the received data and stores it in the update data log 12. The parity update unit 103 calculates a new parity for each data item by using the following formula and updates the database 11.

$$\text{New parity} =$$

$$\text{Old parity} \oplus \text{RECEIVED New data} \oplus \text{RECEIVED Old data}$$

where RECEIVED is an exclusive logical sum. There is no old data for a record addition, and there is no new data for a record deletion.

The parity update unit 103 of the computer collects periodically or at an auditing time the latest data in each database of each of the data bases 21, 31 and 41 of the computers 2, 3 and 4, generates a parity for each data item of each record of each database, and compares this parity with the corresponding parity stored in the database 11. If both the parities are not coincident, it means that fraudulent data update was performed at some branch. Examples of such fraudulent data update are not reporting report data update by the computer at some branch to the computer 1 at the main office, and reporting update data different from the update data stored in the database at some branch to the computer 1 at the main office. In the latter case, the parity update unit 103 detects the database name, row number, and data item for which the fraudulent data update was performed, and thereafter, in accordance with the detected information, an unrepresented processing unit of the computer 1 searches the update data log 12 to identify the branch name and reception time when the data item was updated fraudulently.

In the above embodiment, parities are generated for all data items stored in databases. For the parity comparison, a parity only for a specific data item such as an order number and a money amount may be used.

Fig. 4 is a schematic diagram showing the structure of a computer system according to another embodiment of the invention. This system has an encrypting server 6 in addition to the structure of the system shown in Fig. 1. The processes to be executed by the parity update unit 103 and log generator unit 101 are modified to change the contents of the update data log 12. The server 6 adds a current time to data received from the computers 2, 3 and 4, encrypts the data added with the current time and transmits it to the computers 1, 2, 3 and 4. The parity update unit 103 decrypts the encrypted data received from the server 6 via the network 5, generates a parity for the data after and before update to store it in the database 11, and operates to store a time when the parity was generated in the update data log 12. In this embodiment, a digital signature system using a pu-

bic key is used in which data is enciphered by using a signature key (secret key) and deciphered by using a public key.

Fig. 5A is a flow chart illustrating the procedures to be executed by the computers 2, 3 and 4, server 6 and computer 1. As operators enter update data, the computers 2, 3 and 4 at the branches update the databases 21, 31 and 41 (Step 1000), and send data before and after update to the server 6 (step 1010). Upon reception of these data, the server 6 adds a current time to the data (Step 1020), encrypts the whole of the data and current time with the signature key (Step 1030), and transmits it to the data transmitted computer 2, 3 or 4 and to the monitor computer 1 via the network 5 (Step 1040). The parity update unit 103 of the computer 1 receives the encrypted data via the communication unit 102 and decrypts it with the public key (step 1050). New and old data are picked up from the decrypted data (Step 1060) to generate a new parity by using the above-described formula (Step 1070). Next, the parity update unit 103 instructs the server 6 to send the current time (Step 1080). As the server 6 sends the current time (Step 1090), the parity update unit 103 received the current time sends it to the log generator unit 101. The log generator unit 101 decrypts the data received from the server 6 via the communication unit 102 to generate log information which is added with the current time received from the parity update unit 103 and stored in the update data log 12 (Sep 1100). The parity update unit 103 replaces the corresponding old parity in the monitor database 11 by the new parity. The computers 2, 3 and 4 at branches receive the encrypted data from the server 6 and store it as auditing submission data (Step 1110). Figs. 5A to 5D show operation steps (programs) of the branch computers, server and center computer. These operation steps are realized by programs to be executed by a corresponding computer, and these programs may be stored in a storage medium including a computer readable memory.

Fig. 6 is a flow chart illustrating the procedures to be executed by the computers 2, 3, 4 and 1 at an auditing time. The computers 2, 3 and 4 transmit the stored submission data to the computer 1 (Step 2000). An unrepresented processing unit of the computer 1 receives the transmitted data and decrypts it with the public key (Step 2010). If the decrypted data does not contain a defined data attribute, it is judged that there was a fraudulent data update at some branch (Step 2020). A record corresponding to the received data is acquired from the update data log 12 and a parity update time is derived from this record (Step 2030) to compare it with the data update time received from the branch (Step 2040). If the parity update time is older than the data update time, it is judged that there was a fraudulent data update at the branch (Step 2040). Programs for realizing branch audit process steps may be stored in a storage medium including a computer readable memory. Programs for realizing the process steps shown in Figs. 5B to 5D and

Fig. 6 may be downloaded from an external system into a corresponding computer.

In this embodiment, if the computers 2, 3 or 4 transmits to the computer 1 the data which was not processed by the server 6, the computer cannot decrypt this data so that fraudulent data update can be detected. Furthermore, if the parity update time is older than the data update time for the same data item, it can be determined that the data was updated fraudulently.

According to the present invention, the contents of a database are converted into parity data for the data comparison and verification. Therefore, fraudulent data update by insider frauds can be detected. The invention is effectively applicable to on-line institutes such as banking systems and stock exchanging systems.

### Claims

1. A method of detecting fraudulent data update made by distributed computers having a plurality of databases with a common data format, comprising the steps, in a monitor computer for monitoring data update of said plurality of databases, of:

generating parity data of initial data stored in the databases at each of same storage fields and storing the generated parity data;  
generating, each time data in each database is updated, new parity data from data before and after the update and old parity data to replace the old parity data by the new parity data;  
comparing parity data generated at a specific time from latest data stored in the databases at each of the same storage fields, with the parity data already stored; and  
determining, if the comparison result of said comparing step indicates an inconsistency of both the parity data, that data in the databases has been updated fraudulently.

2. A method according to claim 1, further comprising the steps, to be executed by the monitor computer, of:

storing log information in a storage unit of the monitor computer, said log information including an identifier of a data updated computer, a data updated time, and data before and after the update; and  
searching, when a fraudulent data update is detected, corresponding data from the log information, and deriving information including the identifier of the data updated computer and the data updated time out of the corresponding data.

3. A method according to claim 1, wherein said mon-

itor computer executes the steps of:

generating parity data for a specific data item  
preselected from the databases and storing the  
generated parity data; and  
comparing the parity data for the specific data  
item with the corresponding parity data gener-  
ated at the specific time.

4. A method according to claim 1, wherein update data transmitted from the distributed computers is encrypted and then transmitted to the monitor computer, and the monitor computer checks whether the received update data can be decrypted, and if the update data cannot be decrypted, judges that the update data is fraudulent.
5. A storage which stores parity data to be used for detecting fraudulent data update made by distributed computers having a plurality of databases with the common data format, wherein a monitor computer for monitoring data update of the plurality of databases generates parity data of initial data stored in the databases at each of same storage fields and storing the generated parity data and generates, each time data in each database is updated, new parity data from data before and after the update and old parity data to replace the old parity data by the new parity data.
6. A system for monitoring fraudulent update of distributed databases, said system configuring an intranet including a plurality of distributed databases (21, 31, 41) having the common data format, distributed computers (2, 3, 4) connected to respective ones of the distributed databases, a network (5) connected to the distributed computers, and a monitor computer (1) connected to the network for monitoring data update of the distributed databases, wherein said monitor computer comprises:

a monitor database (11) connected to the monitor computer;  
means for generating parity data of initial data stored in the databases at each of same storage fields and storing the generated parity data in said monitor database;  
means for generating, each time data in each database is updated, new parity data from data before and after the update and old parity data to replace the old parity data stored in said monitor database by the new parity data;  
means for comparing parity data generated at a specific time from latest data stored in the databases at each of the same storage fields, with the parity data stored in said monitor database; and  
means responsive to the comparison result

from said comparing means indicating an inconsistency of both the parity data for determining that data in the databases has been updated fraudulently.

7. A monitor system according to claim 6, wherein said monitor computer is a computer selected from the plurality of distributed computers connected to the network.
8. A monitor system according to claim 6, wherein the databases are relational databases, said initial parity data generating means generates parity data from initial data stored in the relational databases at each of same rows or records, and said comparing and determining means start operating at an auditing time.
9. A monitor system according to claim 6, wherein said monitor computer further comprises:  
  
a storage unit (12) for storing log information including an identifier of a data updated computer, a data updated time, and data before and after the update; and  
means for searching, when said determining means detects fraudulent data update, corresponding data from the log information stored in said storage unit, and deriving information including the identifier of the data updated computer and the data updated time out of the corresponding data.
10. A monitor system according to claim 6, further comprising an encrypting server (6) connected to the network, wherein said server encrypts update data transmitted from the distributed computers and then transmitting the encrypted update data to the monitor computer, and the monitor computer checks whether the received update data can be decrypted, and if the update data cannot be decrypted, judges that the update data is fraudulent.
11. A computer readable recording medium which stores a program to be executed by a monitor computer for detecting fraudulent data update made by distributed computers having a plurality of databases with a common data format, by using a monitor database and a monitor data log, said program to be executed by the monitor computer comprising:

program means for generating parity data of initial data stored in the databases at each of same storage fields and storing the generated parity data;  
program means for generating, each time data in each database is updated, new parity data from data before and after the update and old

parity data to replace the old parity data stored in the monitor database by the new parity data; program means for comparing parity data generated at a specific time from latest data stored in the databases at each of the same storage fields, with the parity data stored in the monitor database; and  
 program means responsive to the comparison result indicating an inconsistency of both the parity data for determining that data in the databases was updated fraudulently.

means each load the program stored in the recording medium recited in claim 11 in a memory of the monitor computer and operate by executing a corresponding program part by the monitor computer.

12. A recording medium according to claim 11, wherein the program to be executed by the monitor computer further comprises:

program means for storing log information in a storage unit of the monitor computer, the log information including an identifier of a data updated computer, a data updated time, and data before and after the update; and  
 program means for searching, when a fraudulent data update is detected, corresponding data from the log information, and deriving information including the identifier of the data updated computer and the data updated time out of the corresponding data.

13. A recording medium according to claim 11, wherein said program to be executed by the monitor computer comprises:

program means for generating parity data for a specific data item preselected from the databases and storing the generated parity data; and  
 program means for comparing the parity data for the specific data item with the corresponding parity data generated at the specific time.

14. A recording medium according to claim 11, wherein said program to be executed by the monitor computer comprises:

program means for encrypting update data transmitted from the distributed computers and then transmitting the encrypted update data to the monitor computer; and  
 program means for checking whether the received update data can be decrypted, and if the update data cannot be decrypted, judging that the update data is fraudulent.

15. A system for monitoring fraudulent update according to claim 6, wherein said monitor database, said initial parity data generating and storing means, said parity data update means, said parity data comparing means, and said fraud determining

FIG. 1

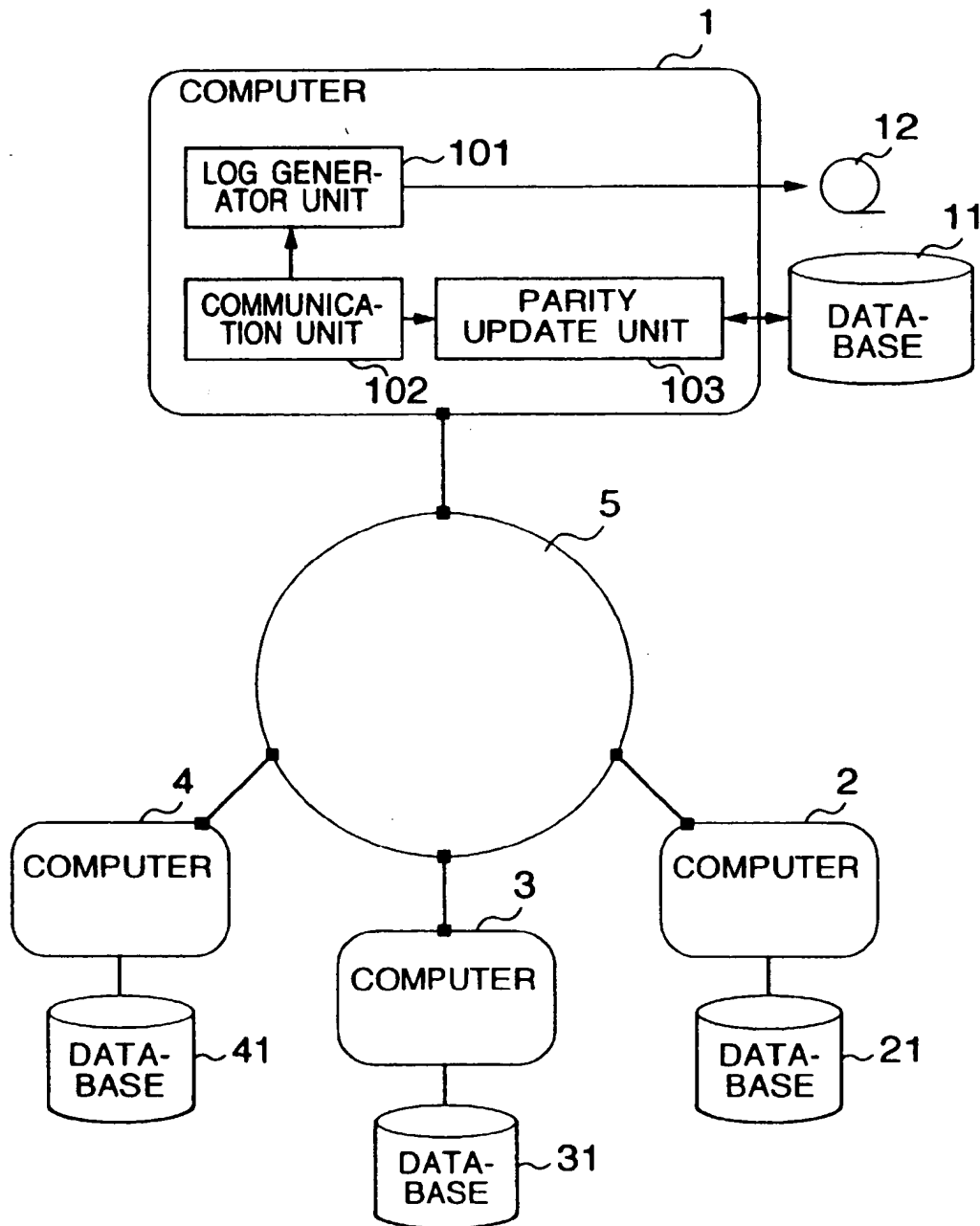


FIG. 2

21, 31, 41

ROW NO.	DATE (M / D)	CUSTOMER NO.	PRODUCT NO.	QUANTITY	MONEY AMOUNT
1	01/20	0125	573	10	20000
2					
3					
4					
5					
⋮					
⋮					
⋮					
n					

FIG. 3

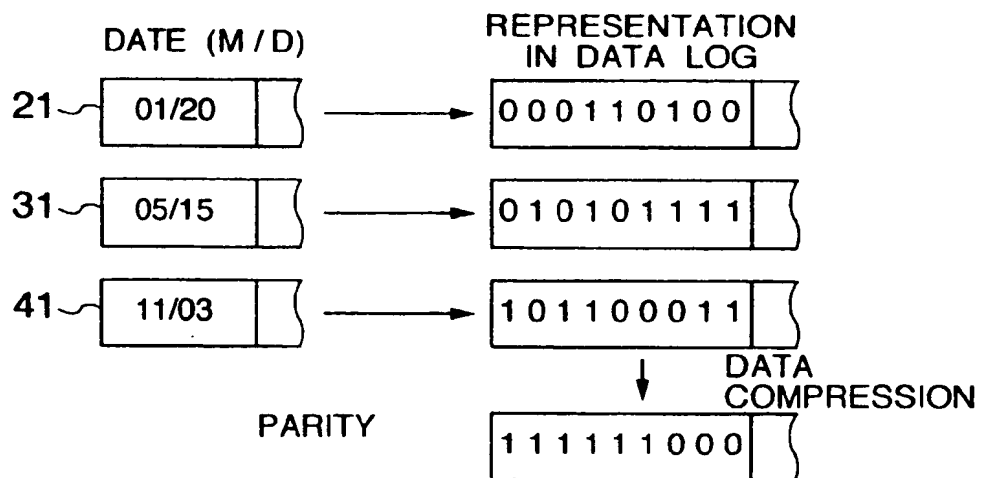


FIG. 4

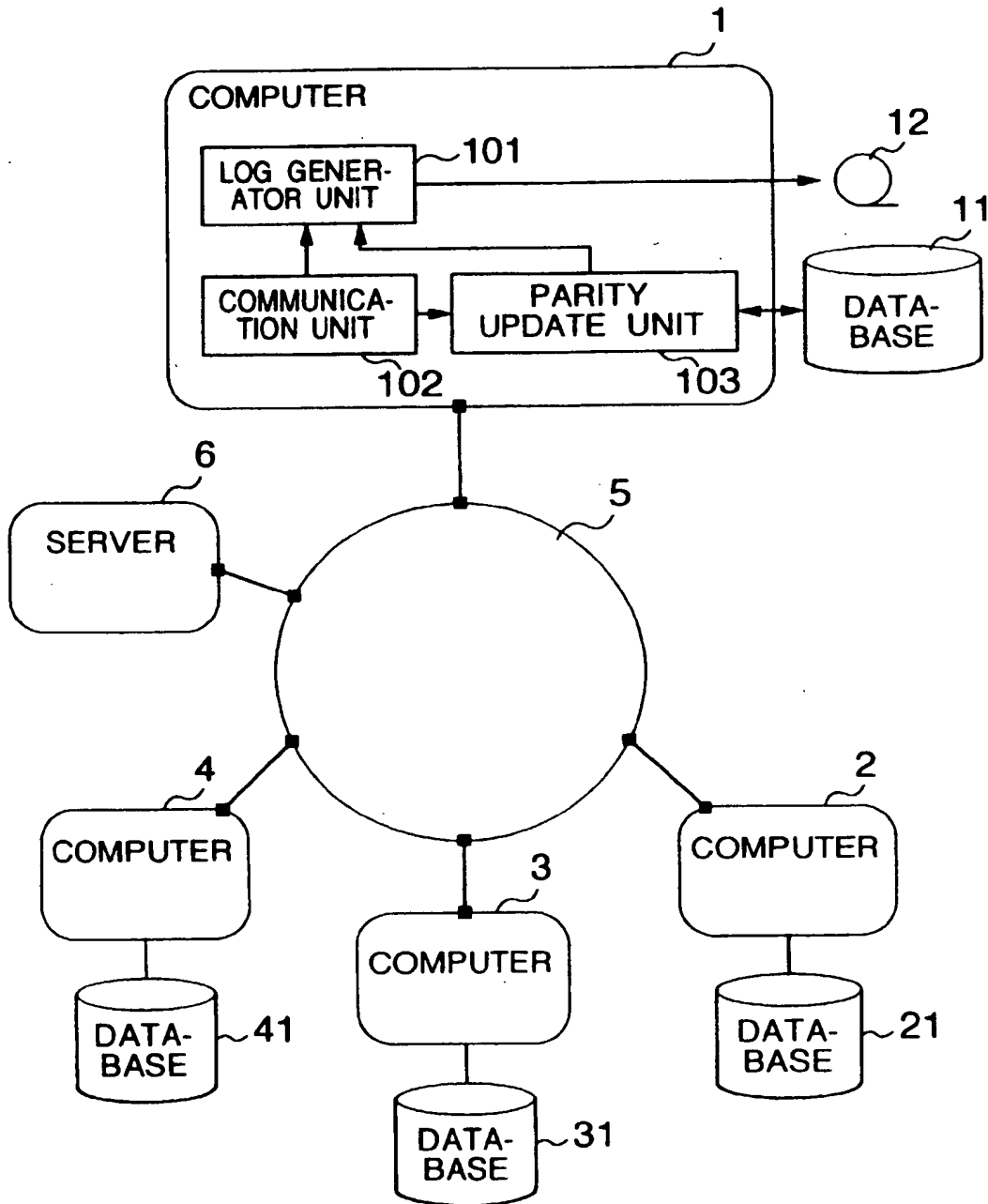


FIG. 5A

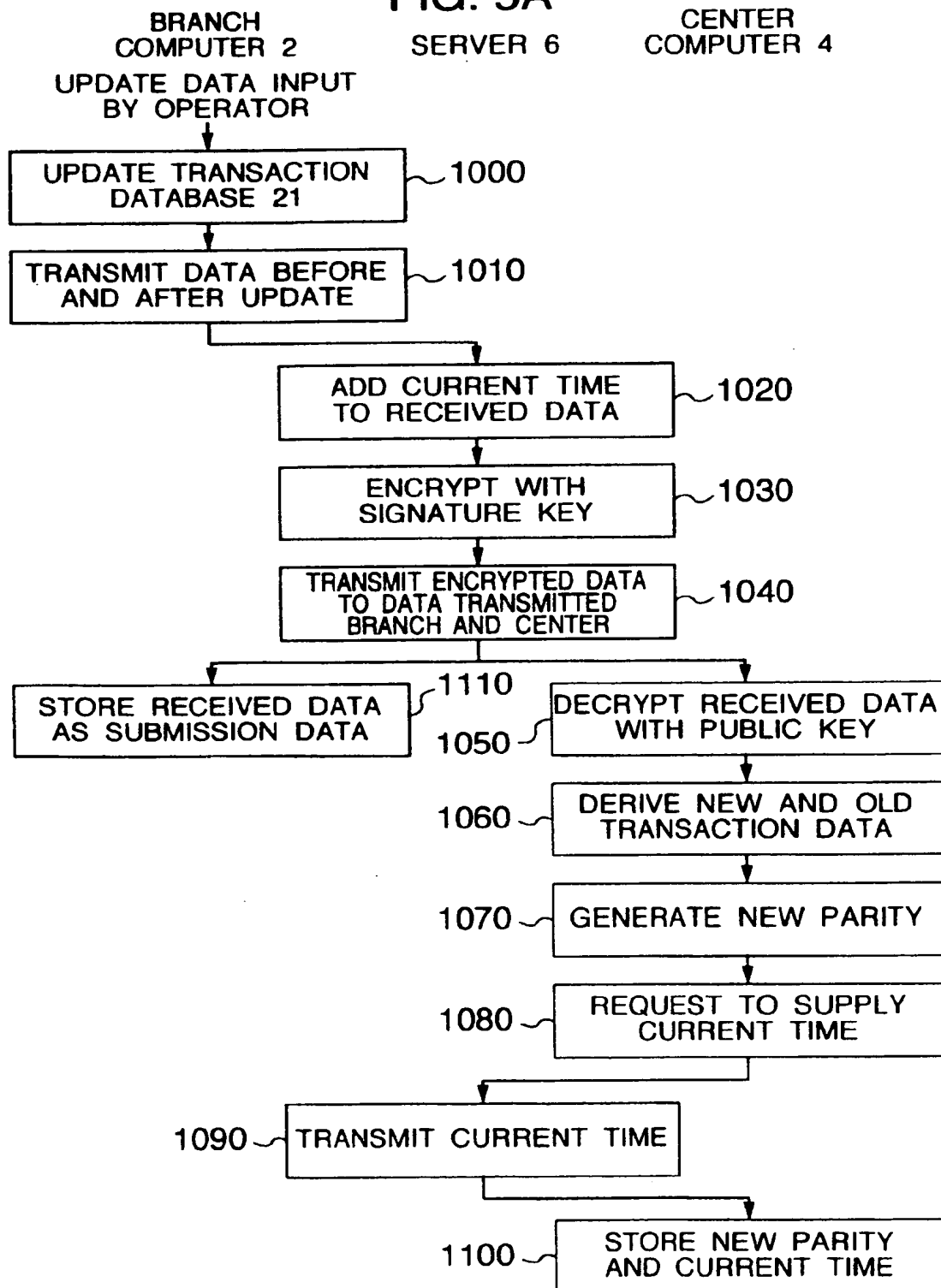


FIG. 5B

BRANCH COMPUTERS 2-4

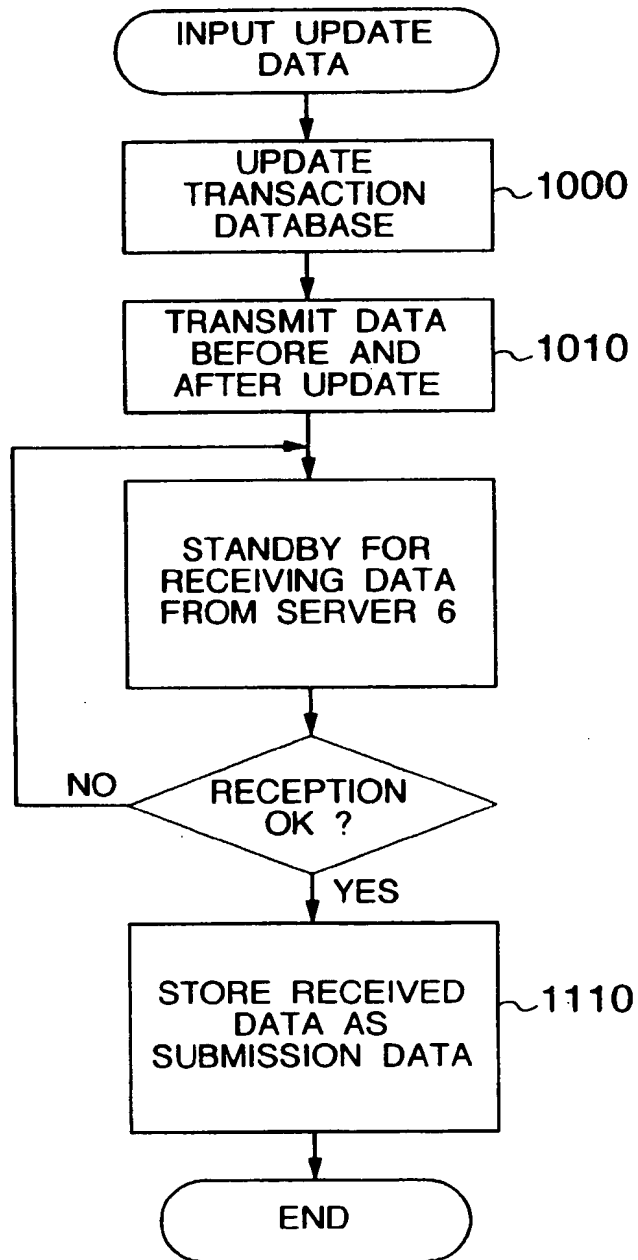


FIG. 5C

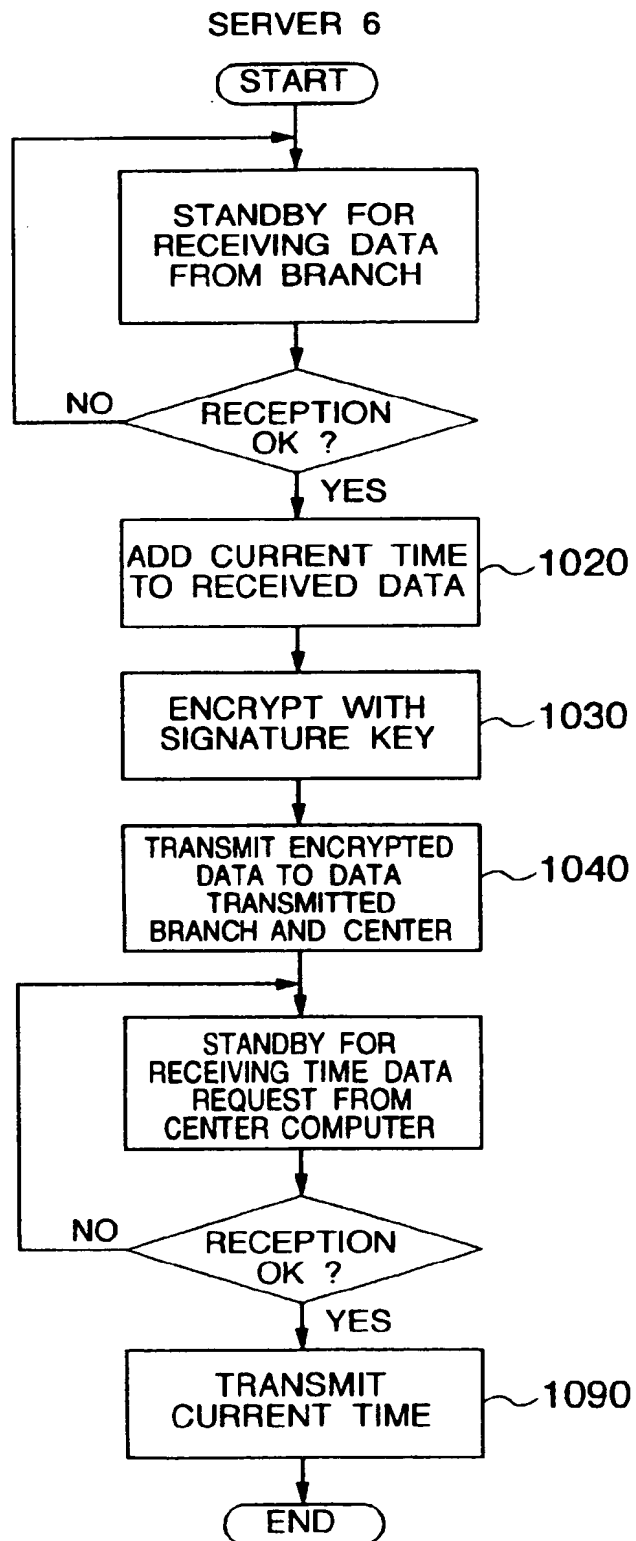


FIG. 5D

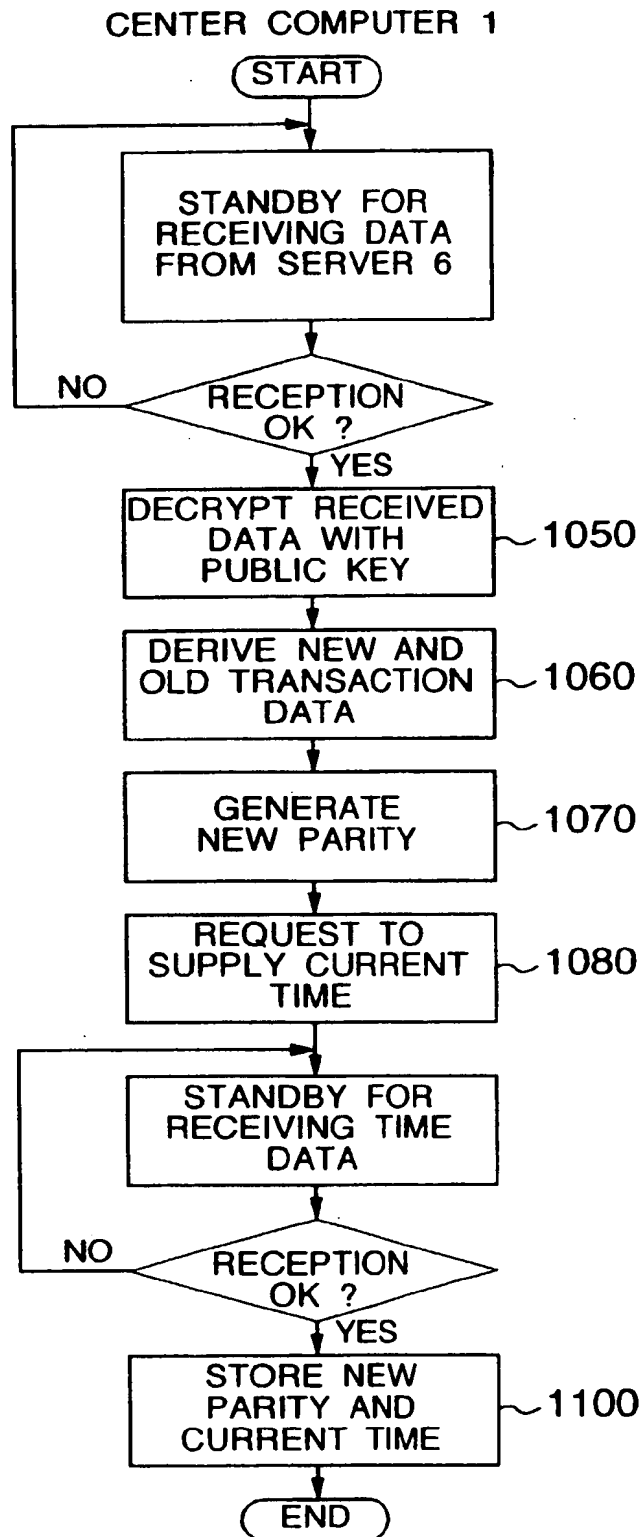
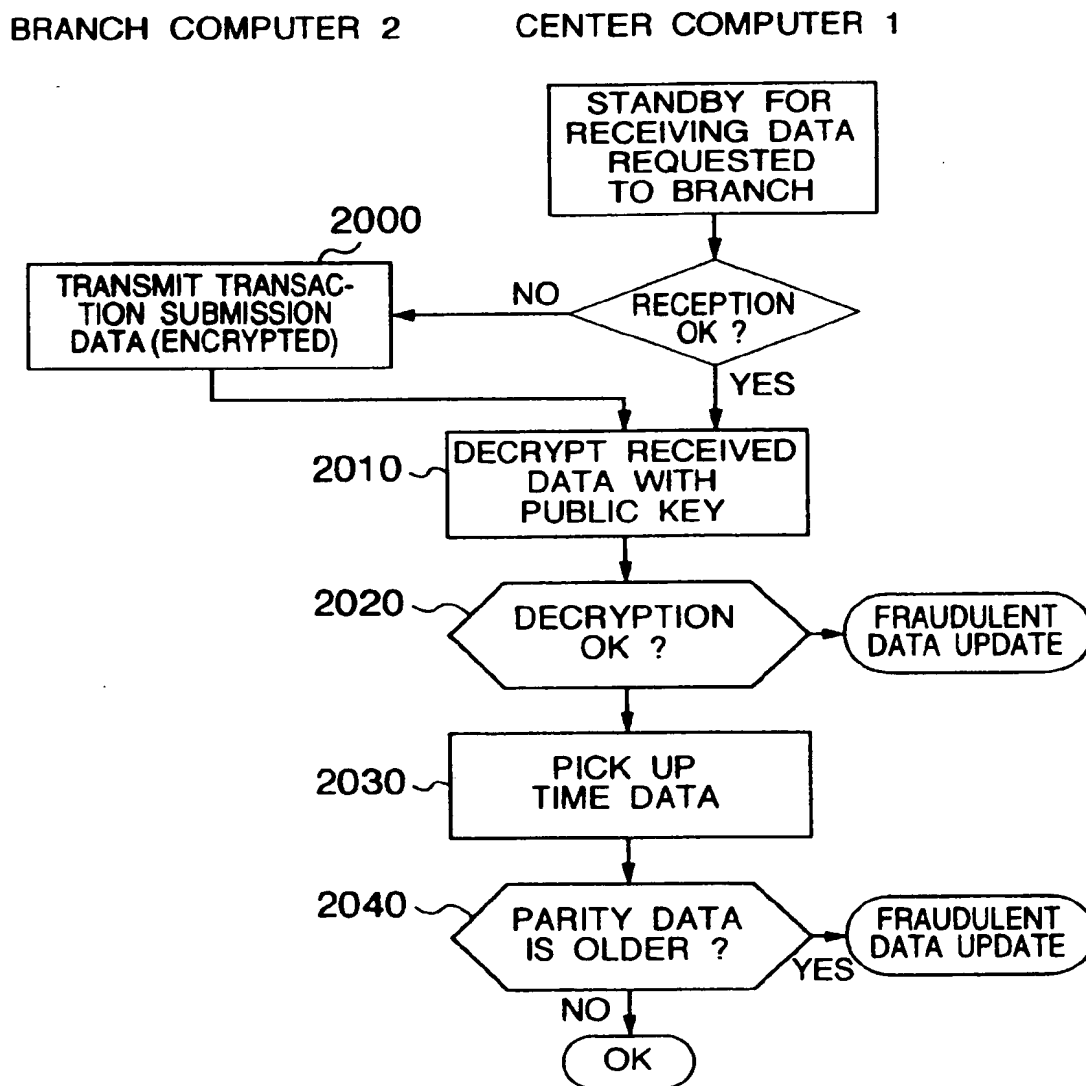


FIG. 6





European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number  
EP 97 30 3552

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
A	EP 0 654 920 A (FISCHER ADDISON M) 24 May 1995 * column 2, line 54 - column 3, line 4 *	1-15	G06F1/00
A	EP 0 635 957 A (DEUTSCHE BUNDESPOST TELEKOM) 25 January 1995 * column 1, line 1-48; claim 1 *	1-15	
A	COMPUTERS & SECURITY INTERNATIONAL JOURNAL DEVOTED TO THE STUDY OF TECHNICAL AND FINANCIAL ASPECTS OF COMPUTER SECURITY, vol. 13, no. 7, 1 January 1994, pages 573-580, XP000485118 HARDJONO T ET AL: "DATABASE AUTHENTICATION REVISITED" * page 574, left-hand column, line 21-27 *	3,13	
A	SIGSAC REVIEW, vol. 10, no. 2 / 03, 1 January 1992, pages 44-62, XP000358274 ALBERT S ET AL: "REFERENCE MODEL FOR DATA MANAGEMENT SECURITY AND PRIVACY" * page 57, paragraph 3.2.2 *	4,6,14	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
			G06F H04L
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 20 August 1997	Examiner Huyghe, E
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date U : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

EP 0 814 398 A1 (P) (M) (C) (I)

**This Page Blank (uspto)**